

Clustering Data Nilai Mahasiswa Untuk Pengelompokan Konsentrasi Jurusan Menggunakan Fuzzy Cluster Means

Tb. Ai Munandar

Teknik Informatika, Fakultas
Teknologi Informasi
Universitas Serang Raya
Banten, Indonesia
e-mail : tebe_aja@yahoo.com

Wahyu Oktri Widyarto

Teknik Industri, Fakultas Teknik
Universitas Serang Raya
Banten, Indonesia
e-mail : woktri_ok@yahoo.com

Harsiti

Sistem Informasi, Fakultas
Teknologi Informasi
Universitas Serang Raya
Banten, Indonesia
e-mail : harsiti@yahoo.com

Abstrak—Menentukan konsentrasi jurusan adalah hal sulit dan sering mendatangkan permasalahan baru di masa mendatang. Sering terjadinya ketidakcocokan mahasiswa terhadap konsentrasi yang saat ini dipilihnya, mengakibatkan tidak fokusnya proses pembelajaran mahasiswa karena kesulitan untuk menyerap materi sesuai dengan konsentrasi yang dipilihnya. Pada penelitian ini dibahas mengenai implementasi metode *fuzzy cluster-means* terhadap data-data bobot nilai mahasiswa pada matakuliah tertentu untuk memperoleh pengelompokan konsentrasi jurusan menurut perolehan nilai akademik berdasarkan kemiripan yang dimilikinya. Analisa data dilakukan dengan bantuan aplikasi MATLAB untuk pembentukan cluster data sesuai dengan yang diharapkan. Hasil penelitian ini berupa tiga buah data cluster yang bisa digunakan untuk pendukung keputusan terhadap penentuan konsentrasi dan dari 126 data mahasiswa, diperoleh sebanyak 28 mahasiswa masuk ke dalam cluster 1(multimedia), 70 mahasiswa pada cluster 2 (web) dan 28 mahasiswa pada cluster 3 (pemrograman).

Kata kunci—konsentrasi jurusan, *fuzzy-cluster-means*, bobot nilai, MATLAB, data cluster

I. PENDAHULUAN

Penentuan konsentrasi jurusan bagi mahasiswa tidak mudah untuk dilakukan. Penentuan konsentrasi akan membantu mahasiswa lebih fokus terhadap apa yang diminati dan disesuaikan dengan nilai akademis yang dimilikinya [1]. Program Studi Teknik Informatika Universitas Serang Raya, memiliki tiga konsentrasi jurusan yang bisa dipilih, yakni konsentrasi Pemrograman, Web dan Multimedia. Saat ini, penentuan konsentrasi jurusan yang dilakukan, hanya berlandaskan keinginan kuat dari mahasiswa ketika akan dilakukan pengkonsentrasian jurusan. Tidak adanya sistem atau aturan yang baku dalam penentuan konsentrasi jurusan sering membawa dampak signifikan terhadap kegiatan belajar mahasiswa serta ketidak mampuan mahasiswa menentukan topik tugas akhir apa nanti yang akan dibuat. Lebih jauh lagi, banyak mahasiswa merasa salah mengambil konsentrasi di tengah semester yang sedang berjalan sehingga mengakibatkan terjadinya pertukaran kelas dan konsentrasi yang tidak semestinya. Penentuan konsentrasi terlalu dini, juga merupakan masalah lain bagi mahasiswa, karena mata kuliah dasar untuk

konsentrasi jurusan belum sepenuhnya terserap secara mendalam.

Tugas besar bagi institusi pendidikan untuk mampu mengenali kemampuan akademis mahasiswanya dan mengarahkannya memasuki konsentrasi jurusan yang tepat. Banyak metode yang bisa digunakan untuk melakukan penentuan konsentrasi jurusan, salah satunya adalah dengan menggunakan konsep *Fuzzy-Cluster-Means* (FCM) yakni dengan cara mengelompokan data-data nilai mahasiswa menurut kemiripan yang dimilikinya. Metode ini dipilih karena kemampuannya untuk melakukan pengelompokan suatu objek/data yang belum memiliki klasifikasi, ke dalam kelas tertentu menurut kesamaan yang dimilikinya [2],[3] berdasarkan derajat keanggotaan dengan cara minimalisasi nilai fungsi objektifnya [4],[5].

Penelitian ini bertujuan untuk menganalisa data mahasiswa menurut bobot nilai mata kuliah tertentu menggunakan konsep FCM, sehingga mampu memberikan sebuah pendukung keputusan tentang pengelompokan data nilai mahasiswa untuk menentukan klasifikasi terhadap konsentrasi yang mana, seorang mahasiswa seharusnya dimasukan. Hasil penelitian diharapkan mampu menjadi salah satu acuan untuk menentukan konsentrasi jurusan mahasiswa dengan melihat batas bobot nilai akademis mata kuliah tertentu sesuai dengan hasil pengelompokan yang dilakukan menggunakan FCM, tanpa mengesampingkan minat personal mahasiswa

II. TINJAUAN PUSTAKA

Fuzzy-Cluster-Means telah banyak digunakan sebagai alat bantu pendukung keputusan untuk berbagai keperluan. Beberapa diantaranya, digunakan pada bidang *digital image processing* [6] baik penggunaan FCM murni maupun gabungan dengan metode lainnya [7], seperti pendeteksian area tumor otak dengan melakukan segmentasi pada MRI *image* otak [8], menganalisa gambar mikroskopis sebuah sel untuk menganalisa perubahan sell dengan bantuan komputasi morfologis [9], klasifikasi *Synthetic Aperture Radar* (SAR) *image* [10], analisa perubahan bentuk lipatan sclerosis untuk mendeteksi kelainan pada sistem syaraf pusat [11], sistem pendeteksi gangguan pada jaringan komputer [12], pendeteksian area kulit manusia menggunakan media *image*,

yang digabungkan dengan metode segmentasi *hill climbing* dan *bayes rule* [13] dan bagaimana metode level set digunakan untuk melakukan segmentasi *image* [14]

III. FUZZY CLUSTER MEANS (FCM)

Fuzzy Cluster Means (FCM) merupakan algoritma yang digunakan untuk melakukan clustering data berdasarkan keberadaan tiap-tiap titik data sesuai dengan derajat keanggotaannya [15]. Berikut ini adalah algoritma *clustering* FCM [4] :

1. *Input* data yang akan di-cluster X, berupa matrik berukuran $n \times m$ (n = jumlah *sample* data, m = atribut setiap data), dimana X_{ij} adalah data *sample* ke- i ($i=1,2,3,...,n$), atribut ke- j ($j=1,2,...,m$).
2. Selanjutnya, tentukan nilai-nilai awal perhitungan seperti, jumlah *cluster*, pangkat, iterasi maksimal, eror terkecil yang diharapkan (ξ), fungsi objektif awal dan iterasi awalnya.
3. Bangkitkan nilai random dalam bentuk elemen matriks partisi awal U (μ_{ik} , $i=1,2,...,n$; $k=1,2,...,c$), kemudian hitunglah jumlah setiap kolom dengan persamaan sebagai berikut :

$$Q_i = \sum_{k=1}^c \mu_{ik} \quad (1)$$

Selanjutnya tentukan nilai matriks partisi awal, dengan persamaan sebagai berikut :

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (2)$$

4. Hitung pusat *cluster* ke- k ; V_{kj} , dimana $k=1,2,...,c$; dan $j=1,2,...,m$, dengan persamaan sebagai berikut :

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (3)$$

5. Selanjutnya, lakukan perhitungan nilai fungsi objektif pada iterasi ke- t (P_t), dengan persamaan berikut :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} * V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (4)$$

6. Hitung perubahan matriks partisi U, dengan persamaan sebagai berikut :

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-\frac{1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-\frac{1}{w-1}}} \quad (5)$$

Dimana $i=1,2,...,n$; dan $k=1,2,...,c$;

7. Langkah terakhir adalah dengan mengecek kondisi berhenti, dengan ketentuan sebagai berikut
 - a. Jika $(|P_t - P_{t-1}| < \xi$ atau $t >$ iterasi maksimal) maka berhenti
8. Jika tidak, maka $t=t+1$ kemudian ulangi langkah ke-4.

IV. METODOLOGI PENELITIAN

Penelitian ini diawali dengan pengumpulan data bobot nilai mata kuliah tertentu yang dimiliki oleh masing-masing mahasiswa. Variabel yang terlibat sebanyak sepuluh variabel yang terdiri atas variabel X1 (matakuliah Pemrograman Internet dan HTML), X2 (matakuliah Algoritma dan Pemrograman), X3 (matakuliah Pemrograman Terstruktur), X4 (matakuliah Komunikasi Data), X5 (matakuliah Struktur Data), X6 (matakuliah Matematika Diskrit), X7 (nilai IP Semester 1), X8 (nilai IP Semester 2), X9 (nilai IP Semester 3) dan X10 (nilai IP Semester 4). Data bobot nilai dari masing-masing variabel di atas kemudian diolah menggunakan FCM untuk menentukan data *clustering* yang nantinya akan digunakan untuk penentuan konsentrasi jurusan.

Tools FCM dari aplikasi MATLAB digunakan untuk melakukan analisa dan peng-cluster-an data sesuai dengan jumlah yang diharapkan. Setelah data pengclusteran diperoleh, semua data awal kemudian diuji ulang menggunakan hasil *clustering* untuk mengetahui setiap mahasiswa masuk ke dalam kelompok data nilai yang mana.

V. HASIL DAN PEMBAHASAN

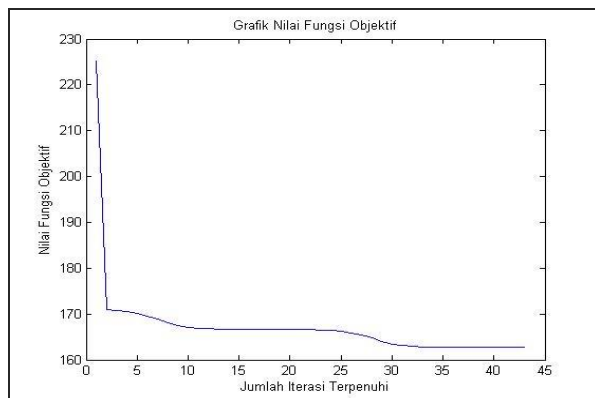
Penelitian ini menguji sebanyak 126 data bobot nilai mahasiswa sesuai dengan variabel yang telah ditentukan (lihat Tabel I). Data pengujian terdiri atas sepuluh variabel, di mana, enam variabel merupakan variabel untuk bobot nilai mata kuliah, dan empat variabel lainnya merupakan nilai Indeks Prestasi Mahasiswa dari semester satu sampai empat. Bobot nilai yang dijadikan data pengujian merupakan bobot nilai mutu terhadap nilai huruf yang terbagi ke dalam 8 kategori nilai mutu, yaitu A = 4.00; A- = 3.75; B+ = 3.50; B = 3.00; B- = 2.75; C+ = 2.50; C = 2.00; D = 1.00.

TABEL I. SAMPEL DATA BOBOT NILAI MAHASISWA

Mhs Ke-	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	1,00	4,00	2,00	2,00	3,50	3,75	2,09	2,70	3,15	2,85
2	4,00	4,00	3,00	3,75	3,50	3,75	3,55	3,36	3,65	3,30
3	4,00	4,00	4,00	3,75	3,50	3,75	3,74	3,66	3,68	3,74
4	3,50	2,00	3,00	3,00	3,50	3,50	3,04	3,36	3,48	3,15
5	4,00	3,00	4,00	3,50	3,50	3,00	3,27	3,57	3,75	3,39
6	3,75	2,00	4,00	3,00	3,50	3,00	3,25	3,48	3,13	3,12
7	4,00	2,50	4,00	3,00	3,50	3,50	3,50	3,17	3,69	3,65
8	3,75	1,00	4,00	3,00	3,50	3,00	3,05	2,72	3,48	3,17
9	3,00	2,50	3,00	3,00	3,50	3,00	2,66	2,78	3,40	3,08
10	4,00	4,00	4,00	4,00	3,50	3,00	3,64	3,71	3,45	3,38
11	3,75	2,75	2,75	4,00	3,50	3,75	3,55	3,42	3,67	3,82
...

Sumber : Kantor Akademik Fakultas Teknologi Informasi setelah diolah

Pengujian menggunakan metode FCM dilakukan dengan bantuan *tools* dari aplikasi MATLAB. Data sebelumnya disimpan ke dalam format *.csv dan dimasukkan ke dalam *tools* FCM pada MATLAB dan dimasukkan kedalam fungsi pengujian FCM, dengan memasukan beberapa parameter awal sebelum *clustering* dilakukan. Jumlah *cluster* yang diinginkan pada kasus ini adalah 3 (tiga) *cluster*, sesuai dengan jumlah konsentrasi yang ada di program studi teknik informatika. Jumlah iterasi maksimum diseting sebanyak 100 (seratus) kali dengan *error* terkecil yang diharapkan sebesar 0.00001, dengan nilai perpangkatan adalah 2, fungsi objektif awal adalah 0 (nol) dan iterasi awal adalah 1 (satu)



Gambar 1. Grafik nilai fungsi objektif

Hasil pengujian menunjukkan bahwa proses peng-*cluster*-an berhenti pada iterasi ke 43 dengan nilai fungsi objektif adalah 162.75 (lihat grafik nilai fungsi objektif pada Gambar 1), dan menghasilkan tiga buah *clustering* seperti yang terlihat pada tabel 2 di bawah ini. Hasil *clustering* ini mengindikasikan proses penentuan konsentrasi jurusan bagi mahasiswa sesuai dengan nilai akademis yang dimilikinya

Berdasarkan ketiga *cluster* yang terbentuk, mahasiswa dapat dikelompokkan ke dalam tiga konsentrasi menurut nilai-nilai yang memenuhi pada masing-masing variabel di setiap *cluster*-nya (lihat Tabel II). Ketiga *cluster* tersebut adalah, *cluster* pertama yang diindasikan sebagai konsentrasi Multimedia, *cluster* kedua sebagai konsentrasi Web dan *cluster* ketiga sebagai konsentrasi Pemrograman. Menurut hasil *cluster*, mahasiswa akan dikategorikan ke dalam *cluster* 1 (asumsi konsentrasi Multimedia) jika memiliki nilai mutu mata kuliah Algoritma dan Pemrograman sebesar 3.01, Pemrograman Internet dan HTML sebesar 3.67, Pemrograman Terstruktur sebesar 2.01, Komunikasi Data sebesar 2.87, Struktur Data sebesar 3.17, Matematika Diskrit sebesar 2.86, nilai IP Semester 1 sebesar 3.29, nilai IP Semester 2 sebesar 2.71, nilai IP Semester 3 sebesar 3.14 dan nilai IP Semester 4 sebesar 2.99.

TABEL II. DATA HASIL CLUSTERING YANG TERBENTUK

Clust. Ke-	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	3,01	3,67	2,01	2,87	3,17	2,86	3,29	2,71	3,14	2,99
2	3,37	2,16	3,08	3,21	3,23	2,97	3,12	2,91	3,37	3,12
3	3,72	3,43	3,55	3,55	3,51	3,17	3,56	3,34	3,53	3,43

Kategori *cluster* 2 (asumsi konsentrasi Web), diindikasikan jika mahasiswa memiliki nilai mutu mata kuliah Algoritma dan Pemrograman sebesar 3.37, Pemrograman Internet dan HTML sebesar 2.16, Pemrograman Terstruktur sebesar 3.08, Komunikasi Data sebesar 3.21, Struktur Data sebesar 3.23, Matematika Diskrit sebesar 2.97, nilai IP Semester 1 sebesar 3.12, nilai IP Semester 2 sebesar 2.91, nilai IP Semester 3 sebesar 3.37 dan nilai IP Semester 4 sebesar 3.12.

TABEL III. SAMPLE HASIL PENGKONSENTRASIAN MENURUT HASIL CLUSTERING

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	C1	C2	C3
1	3,00	4,00	2,00	3,00	3,50	2,75	3,59	3,02	3,14	3,06		*	
2	3,00	4,00	2,00	3,50	3,50	2,75	3,36	3,04	3,00	3,10		*	
3	3,75	4,00	3,50	3,00	3,50	3,00	3,73	3,61	3,68	3,58			*
4	4,00	4,00	3,50	4,00	3,50	3,50	3,84	3,33	3,24	3,54			*
5	2,50	4,00	1,00	3,00	3,50	3,00	3,18	2,50	3,33	3,11		*	
6	4,00	3,00	2,75	3,00	2,75	3,00	3,14	2,61	3,26	3,20		*	
7	4,00	1,00	3,75	2,75	3,50	3,00	2,70	2,83	3,40	3,13	*		
8	4,00	4,00	4,00	3,75	3,50	3,75	3,84	3,12	3,63	3,60			*
9	3,50	2,50	4,00	2,50	3,50	3,00	3,14	2,87	3,38	3,01		*	
10	3,50	2,50	4,00	3,00	3,50	3,00	3,41	2,74	3,68	3,43		*	
11	3,75	3,00	4,00	3,00	3,50	3,50	3,45	3,13	3,61	3,39			*
12	3,50	2,50	4,00	3,00	3,50	3,00	3,41	3,20	3,61	3,00		*	
13	3,00	4,00	2,75	3,75	3,50	3,00	3,34	3,54	3,73	3,37		*	
14	4,00	2,00	4,00	3,75	3,50	3,50	3,52	3,16	3,71	3,26		*	
15	2,00	4,00	1,00	2,50	3,50	2,75	3,09	1,35	3,38	2,99	*		
...

Kategori *cluster* 3 (asumsi konsentrasi Pemrograman), diindikasikan jika mahasiswa memiliki nilai mutu mata kuliah Algoritma dan Pemrograman sebesar 3.72, Pemrograman Internet dan HTML sebesar 3.43, Pemrograman Terstruktur sebesar 3.55, Komunikasi Data sebesar 3.55, Struktur Data sebesar 3.51, Matematika Diskrit sebesar 3.17, nilai IP Semester 1 sebesar 3.56, nilai IP Semester 2 sebesar 3.34, nilai IP Semester 3 sebesar 3.53 dan nilai IP Semester 4 sebesar 3.43

Hasil *clustering* ini kemudian diujikan kembali ke dalam data primer dan menghasilkan sejumlah kelompok konsentrasi dengan jumlah masing-masing mahasiswa sesuai dengan *cluster* yang ada. Mahasiswa yang ter-*cluster* ke dalam *cluster* 1 (Multimedia) sebanyak 28 mahasiswa, *cluster* 2 (Web) sebanyak 70 mahasiswa dan *cluster* 3 (Pemrograman) sebanyak 28 mahasiswa (lihat sampel hasil penelusuran pada Tabel III)

VI. KESIMPULAN

Penelitian ini berhasil melakukan clustering data menurut bobot nilai mata kuliah dan nilai indeks prestasi mahasiswa untuk menentukan cluster konsentrasi jurusan pada program studi teknik informatika Universitas Serang Raya. Hasil pengujian menunjukkan bahwa dari 126 pengujian data mahasiswa, sebanyak 28 mahasiswa masuk ke dalam cluster 1 sehingga dianjurkan untuk ditempatkan pada konsentrasi Multimedia, 70 mahasiswa masuk ke dalam cluster 2 sehingga dianjurkan untuk ditempatkan pada konsentrasi Web dan 28 mahasiswa masuk ke dalam cluster 3 sehingga dianjurkan untuk ditempatkan pada konsentrasi Pemrograman.

DAFTAR PUSTAKA

- [1]. Rakhman, Zahri, Arkham., Wulandari, Nisa, Helmanatun., Maheswara, Geralvin., Kusumadewi, Sri., 2012, "Fuzzy Inference System Dengan Metode Tsukamoto Sebagai Pemberi Saran Pemilihan Konsentrasi (Studi Kasus Jurusan Teknik Informatika UII)", Proceeding Seminar Nasional Teknologi Informasi 2012.
- [2]. Luthfi, Taufiq, Emha., 2007, "Fuzzy C-Means Untuk Clustering Data (Study Kasus : Data Performance Mengajar Dosen)", Proceeding Seminar Nasional Teknologi 2007 (SNT 2007), pp. D1-D7.
- [3]. El-Zaghmouri, Bassam M., Abu-Zanona, Marwan A., 2012, "Fuzzy C-Mean Clustering Algorithm Modification and Adaptation for Applications", World of Computer Science and Information Technology Journal (WCSIT), Vol. 2, No. 1, pp. 42-45
- [4]. Kusumadewi, Sri., Purnomo, Hari., 2010, "Aplikasi Logika Fuzzy Untuk Pendukung Keputusan", Edisi Kedua, Penerbit GRAHA ILMU : Yogyakarta
- [5]. Jain, Neha., Shukla, Seema., 2012, "Fuzzy Databases Using Extended Fuzzy C-Means Clustering", International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 3, pp.1444-1451.
- [6]. Valvi, Piyus., Shah, Brijesh., Shah, Satish., 2012, "Improved fuzzy c-mean algorithm for Medical Image segmentation", International Journal of Engineering Research & Technology (IJERT), Vol. 1, Issue 3, pp. 1 – 3.
- [7]. Nyma, Alamgir., Kang, Myeongsu., Kwon, Yun-Keun., 2012, "A Hybrid Technique for Medical Image Segmentation", Journal of Biomedicine and Biotechnology, Volume 2012, pp. 1 – 7.
- [8]. Bidhe, A.S., Patil, Priyanka., Dhande, Shraddha., 2012, "Brain Segmentation using Fuzzy C means clustering to detect tumour Region", International Journal of Advanced Research in Computer Science and Electronics Engineering, Volume 1, Issue 2, pp. 85 – 90.
- [9]. Nath, C.K., Talukdar, Jyotisma., 2012, "Cell Image Analysis using Fuzzy Clustering and Morphology based computation", International Journal of Computing, Communications and Networking, Volume 1, No.1, pp. 10 – 12.
- [10]. Samanta, Debabrata., Sanyal, Goutam., 2012, "SAR Image Classification Using Fuzzy C-Means", International Journal of Advances in Engineering & Technology, Vol. 4, Issue 2, pp. 508-512.
- [11]. Agboizebeta, Anthony, Imianvan., Chukwuyeni, Jonathan, Obi., 2012, "Cognitive analysis of multiple sclerosis utilizing fuzzy cluster means", International Journal of Artificial Intelligence & Applications (IJAA), Vol.3, No.1, pp. 33 – 45.
- [12]. Narayan, Esh., Singh, Pankaj., Kumar Tak, Gaurav., 2012, "Intrusion Detection System Using Fuzzy C-Means Clustering with Unsupervised Learning via EM Algorithms", VSRD International Journal of CS & IT Vol. 2 (6), pp. 502 – 510.
- [13]. Vijayanandh, R., Balakrishnan, G., 2012, "Hillclimbing Segmentation with Fuzzy C-Means Based Human Skin Region Detection using Bayes Rule", European Journal of Scientific Research, Vol.76, No.1, pp. 95 - 107.
- [14]. Kumar, Sanjay., Kumar Ray, Santosh., Tewari, Peeyush., 2012, "Combined Approach Using Fuzzy Clustering and Local Image Fitting Level Set Method for Global Image Segmentation", Canadian Journal on Image Processing and Computer Vision, Vol. 3, No. 1, pp. 1 – 5.
- [15]. Setiyono, Budi., Isnanto, R. Rizal., 2008, "Pembagian Kelas Kuliah Mahasiswa Menggunakan Algoritma Pengklasteran Fuzzy C-Means", Prosiding Seminar Nasional Teknoin 2008, pp. E15 – E20.